

文章编号:1001-9081(2009)10-2678-03

基于尺度不变特征变换的 Mean-Shift 目标跟踪

杨心力¹, 杨恢先¹, 曾金芳¹, 于洪²

(1. 湘潭大学 材料与光电物理学院, 湖南 湘潭 411105; 2. 琼州学院 物理系, 海南 五指山 572200)

(yangxinli1983@126.com)

摘要:均值漂移(Mean-Shift)目标跟踪算法由于具有快速模板匹配和无参数密度估计等特点,但也存在其固有的缺陷。为了提高该算法的鲁棒性,把目标分成多个区域,对每个区域利用 Mean-Shift 进行跟踪,迭代次数大于 8 的放弃迭代。然后利用尺度不变特征变换(SIFT)剔除那些匹配的关键点数目少的子区域。最后,利用匹配关键点数目多的区域得到目标的位置。实验结果表明该方法在目标受遮挡、尺度变化、旋转、环境场景等变化等具有很强的鲁棒性。

关键词:目标区域划分; 尺度不变特征变换; 均值漂移; 目标跟踪

中图分类号: TP391.41; TP751 **文献标志码:** A

Mean-Shift object tracking based on scale invariant feature transform

YANG Xin-li¹, YANG Hui-xian¹, ZENG Jin-fang¹, YU Hong²

(1. School of Material and Photoelectric Physics, Xiangtan University, Xiangtan Hunan 411105, China;

2. Department of Physics, Qiongzhou University, Wuzhishan Hainan 572200, China)

Abstract: Mean-Shift algorithm performs well in object tracking field because of its advantages of fast pattern matching and non-parametric estimation. However, this algorithm has its inherent deficiencies. In order to improve the robustness of Mean-Shift algorithm, the target was divided into a number of sub-regions in this paper, each sub-region individually used Mean-Shift tracking, and those whose iterations are more than eight times quit. And Scale Invariant Feature Transform (SIFT) was employed to exclude those sub-regions with smaller matching key points. Finally, the object location was obtained according to the sub-regions with more matching key points. Experiments show that the proposed method is of high robustness in situations of occlusion, scale change, rotation, scene change, etc.

Key words: target region division; Scale Invariant Feature Transform (SIFT); Mean-Shift; object tracking

目标跟踪的应用在近年来得到了广泛的关注和长足的发展,针对不同的跟踪环境和运算要求,已经研究和设计出了许多跟踪方法,关于最近几十年目标跟踪算法的进展情况参考文献[1]。文献[2]作者将 Mean-Shift 算法引入了跟踪领域,并成功进行了目标跟踪。由于 Mean-Shift 算法具有快速模板匹配和无参数密度估计等优点,被广泛应用于跟踪领域。但是,由于现实环境的复杂性加上被跟踪目标的遮挡、尺度变化和旋转以及场景变化等,很难保证 Mean-Shift 算法的稳定性。针对 Mean-Shift 跟踪方法以上局限性,本文把目标分成若干子区域,对每个区域分别跟踪,这样即使某个或某些区域受到遮挡其他区域仍可以正确跟踪。利用尺度不变特征变换(Scale Invariant Feature Transform, SIFT)得到各子区域匹配的关键点数量,剔除匹配的关键点数量少的子区域,利用匹配的关键点数量多的子区域确定目标位置。

1 Mean-Shift 目标跟踪

1.1 目标区域划分

Mean-Shift 算法利用颜色直方图信息以及核函数提供的一些空间信息进行跟踪,但是在目标发生遮挡、旋转、尺度变化、场景颜色与目标相近时,目标的颜色直方图信息以及核函数提供的一些空间信息将会受到很大的影响,从而影响跟踪的稳定性。为了提高稳定性把目标区域划分为 4×4 的区域,如图 1 所示,对每个子区域 $R_{i,j}$ ($i = 1, \dots, 4, j = 1, 2, \dots, 4$) 分

别利用 Mean-Shift 算法进行模式匹配。为了便于表达目标子区域用 $R_{i,j}$ 表示,模板子区域用 $R_{i,j}^t$ 表示。

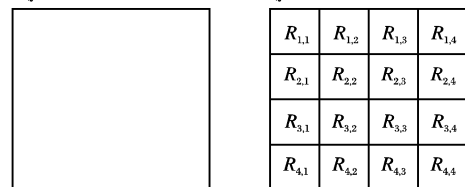


图 1 目标区域分割

1.2 Mean-Shift 目标跟踪算法

1.2.1 目标模型描述

目标模型中心附近的像素比其他像素更可靠,因此对靠近中心的像素赋一个大的权值,而远离中心的像素赋一个小的权值。对目标密度估计引入核函数,则目标模型的概率密度估计可表示为:

$$\hat{q}_u = C \sum_{i=1}^n k \left(\left\| \frac{y_0 - x_i}{h} \right\| \right) \delta [b(x_i) - u] \quad (1)$$

其中, C 是常数, $k(x)$ 为核函数的轮廓函数, 函数 $k(x)$ 中 $\left\| \frac{y_0 - x_i}{h} \right\|$ 的作用是消除不同大小目标的影响。

1.2.2 候选模型描述

在以后的各帧中可能包含运动目标的区域,称为候选区域,其中心坐标为 y , 即核函数的中心坐标。该区域中的像素

收稿日期:2009-04-09。 基金项目:海南省自然科学基金资助项目(60897);海南省教育厅项目(HJ2009-135)。

作者简介:杨心力(1983-),男,河南开封人,助理工程师,硕士研究生,主要研究方向:视频目标跟踪、图像处理; 杨恢先(1964-),男,湖南益阳人,教授,主要研究方向:图像处理、现代信号处理; 曾金芳(1978-),女,湖南娄底人,讲师,主要研究方向:计算机控制、信号处理; 于洪(1972-),男,河北沧州人,副教授,主要研究方向:流媒体技术。

用 $\{x_i\}_{i=1,2,\dots,n_h}$ 表示,对候选区域的描述称为目标候选模型,候选模型的特征值 $u = 1, 2, \dots, m$ 的概率密度是:

$$\hat{p}_u = C_h \sum_{i=1}^{n_k} k\left(\left\|\frac{y-x_i}{h}\right\|^2\right) \delta[b(x_i) - u] \quad (2)$$

其中, C_h 是标准化的常量系数。

1.2.3 相似性函数

在 Mean-Shift 算法中,选 Bhattacharyya 系数作为相似性函数, Bhattacharyya 系数表示如下:

$$B(y) = \sum_{u=1}^m \sqrt{\hat{p}_u(y) \hat{q}_u} \quad (3)$$

$B(y)$ 越大表示两个模型越相似。

目标最可能出现的位置是使 $B(y)$ 最大的候选区域。为使 $B(y)$ 最大,在当前帧中以前一帧中目标中心的位置 y_0 为起始位置开始,寻找最优匹配位置,其中心为 y 。

$$f_{h,k} = \frac{C_h}{2} \sum_{i=1}^{n_k} \omega_i k\left(\left\|\frac{y-x_i}{h}\right\|^2\right) \quad (4)$$

要使 $B(y)$ 最大等价于 $f_{h,k}$ 最大。对式(5)求梯度,经化简可得到:

$$m_{h,c}(x)_1 = \frac{\sum_{i=1}^n x_i w_i g\left(\left\|\frac{x-x_i}{h}\right\|^2\right)}{\sum_{i=1}^n w_i g\left(\left\|\frac{x-x_i}{h}\right\|^2\right)} - x \quad (5)$$

则均值平移向量为:

$$y_{\text{shift}} = \frac{\sum_{i=1}^n x_i w_i g\left(\left\|\frac{x-x_i}{h}\right\|^2\right)}{\sum_{i=1}^n w_i g\left(\left\|\frac{x-x_i}{h}\right\|^2\right)} \quad (6)$$

$$w_i = \sum_{u=1}^m \delta[b(x_i) - u] \sqrt{\frac{\hat{q}_u}{\hat{p}_u}} \quad (7)$$

目标跟踪过程是利用 Mean-Shift 向量梯度下降在当前帧图像中搜索与目标模板特征概率分布相似的目标区域。

2 SIFT 特征匹配

SIFT 算子是一种基于尺度空间,对图像缩放、旋转甚至仿射变换保持不变性的图像局部特征描述算子。SIFT 算法首先在尺度空间进行特征检测,并确定关键点(Keypoints)的位置和关键点所处的尺度,然后使用关键点邻域梯度的主方向作为该点的方向特征,以实现算子对尺度和方向的无关性。

高斯卷积核是实现尺度变换的唯一线性核,于是一幅二维图像的尺度空间定义为:

$$L(x, y, \sigma_L) = G(x, y, \sigma_L) \times I(x, y) \quad (8)$$

其中, $G(x, y, \sigma_L)$ 是尺度可变换的高斯核函数:

$$G(x, y, \sigma_L) = \frac{1}{2\pi\sigma_L^2} \exp[-(x^2 + y^2)/2\sigma_L^2] \quad (9)$$

其中, (x, y) 是坐标, σ_L 是尺度。

输入图像通过不同尺度的高斯核函数滤波和亚采样,形成高斯金字塔图像,然后相邻尺度的两个高斯图像相减得到 DOG (Difference-of-Gaussians) 金字塔多尺度空间。

$$D(x, y, \sigma_L) = L(x, y, k\sigma_L) - L(x, y, \sigma_L) = G(x, y, k\sigma_L) - G(x, y, \sigma_L) \times I(x, y) \quad (10)$$

$$\text{由于: } \frac{\partial G}{\partial \sigma_L} = \sigma_L \Delta^2 G \quad (11)$$

因此:

$$G(x, y, k\sigma_L) - G(x, y, \sigma_L) \approx (k-1)\sigma_L^2 \Delta^2 G \quad (12)$$

对 DOG 尺度空间每个点与相邻尺度和相邻位置的点逐

个进行比较,得到的局部极值位置即为关键点所处的位置和对应的尺度。

在以关键点为中心的邻域窗口内采样,并用直方图统计邻域像素的梯度方向。梯度直方图的范围 $0^\circ \sim 360^\circ$, 其中每 10° 为一个柱,共 36 个柱。直方图的峰值则代表了该关键点的方向,如图 2 所示。后续的描述子构造均以该方向为参照,这样所构造的描述子具有旋转不变在梯度方向直方图中,当存在另一个相当于主峰值的四分之三时,将这个方向定义为关键点的辅方向。一个关键点具有一个主方向和一个以上辅方向,这可以增强匹配的鲁棒性。

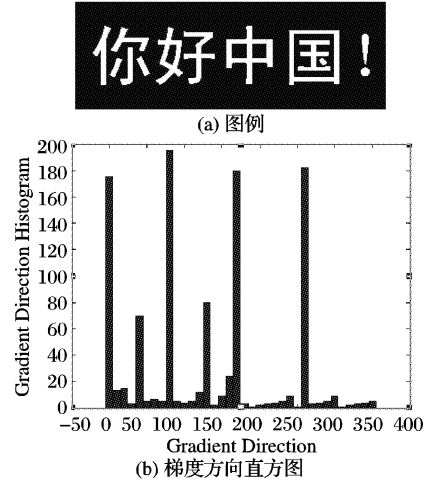


图 2 通过梯度方向直方图确定主方向

以关键点为中心取 8×8 的窗口,在每 4×4 的小块上计算 8 个方向的梯度方向直方图,绘制每个梯度方向的梯度直方图,可形成一个种子点, 2×2 共 4 个种子点,每个种子点有 8 个方向向量信息。这种邻域方向性信息联合的思想增强了算法抗噪声的能力,同时对于含有定位误差的特征匹配也提供了较好的容错性。为了增强匹配的稳健性,对每个关键点可使用 4×4 共 16 个种子点来描述,这样对于一个关键点形成 128 维的 SIFT 特征向量。

当目标子区域 $R_{i,j}$ 的 SIFT 特征向量生成后,采用关键点特征向量的欧氏距离作为目标子区域 $R_{i,j}$ 和模板子区域 $R_{i,j}^T$ 中关键点的相似性判定度量。取模板子区域 $R_{i,j}^T$ 中的某个关键点,其 SIFT 向量表示为 $SIFT_{KP}$ 。在目标子区域 $R_{i,j}$ 中找出其与目标子区域中欧氏距离最近的前两个关键点,它们对应的 SIFT 向量分别表示为 $SIFT_{KP1}$, $SIFT_{KP2}$ 。

$$Dist_1 = \|SIFT_{KP} - SIFT_{KP1}\| \quad (13)$$

$$Dist_2 = \|SIFT_{KP} - SIFT_{KP2}\| \quad (14)$$

假设 $SIFT_{KP1}$ 是最近的距离, $SIFT_{KP2}$ 是次最近距离,则:

$$T = \frac{Dist_1}{Dist_2} \quad (15)$$

如果 $T < \rho$ (本文 $\rho = 0.8$), 关键点匹配, 否则不匹配。如果每个目标子区域 $R_{i,j}$ 与模板子区域 $R_{i,j}^T$ 关键点的个数 $N_{i,j}$ 大于某一阈值 N (本文 $N = 5$) 表示目标子区域 $R_{i,j}$ 与模板子区域 $R_{i,j}^T$ 匹配, 该目标子区域 $R_{i,j}$ 参与目标位置的推算; 否则目标子区域 $R_{i,j}$ 与模板子区域 $R_{i,j}^T$ 不匹配, 该目标子区域 $R_{i,j}$ 不参与目标位置的推算。

假设目标的宽度为 w 高度为 h , 根据匹配目标子区域 $R_{i,j}$ 的位置为 $P'_{i,j}$ 得到的目标的位置为 $P_{i,j}$:

$$P_{i,j} = P'_{i,j} + \begin{pmatrix} (i-2) \times \frac{w}{4} + \frac{w}{8} \\ (j-2) \times \frac{h}{4} + \frac{h}{8} \end{pmatrix} \quad (16)$$

则合成的目标位置 P 为:

$$P = \frac{\sum_{N_{i,j} \geq N} N_{i,j} P_{i,j}}{\sum_{N_{i,j} \geq N} N_{i,j}} \quad (17)$$

式(17)表明匹配的关键点数多,目标子区域 $R_{i,j}$ 与模板子区域 $R_{i,j}^t$ 匹配得好,该位置可靠,赋予大的权值。

3 目标跟踪实验

本实验是在 Windows 环境下利用 VC6.0 和 OpenCV 来进行的。为了检测本文提出方法的稳定性,实验使用一个散打比赛的视频,分别采用传统的 Mean-Shift 算法和本文提出的算法来跟踪目标。当目标受到严重遮挡、迅速旋转和场景的某些景物与目标颜色相近时,传统的 Mean-Shift 算法跟踪失败,而本文提出的算法能很好地跟踪目标。

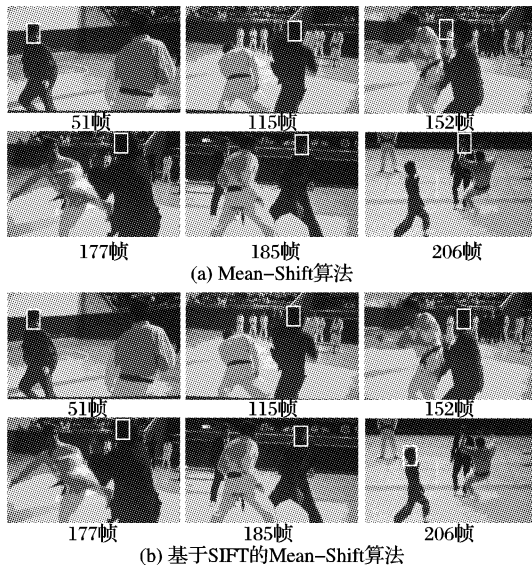


图 3 跟踪实验

跟踪结果如图 3 所示,图 3(a)所示利用传统的 Mean-Shift 跟踪目标,当目标严重遮挡时受到干扰,丢失目标;图 3

(b)所示为基于 SIFT 的 Mean-Shift 目标跟踪,在整个序列能很好地跟踪目标。

4 结语

本文提出的方法把目标区域分成 4×4 的子区域,在目标区域受到遮挡、目标与背景混淆等复杂环境下该方法能实时稳定地跟踪目标,很好解决了目标受到干扰的情况下 Mean-Shift 算法跟踪不稳定的问题。

参考文献:

- [1] YILMAZ A. Object tracking: A survey[J]. ACM Computing Surveys, 2006, 38(4): 13.
- [2] COMANICIU D, RAMESH V, MEER P. Real-time tracking of non-rigid objects using mean-shift[C]// IEEE Conference on Computer Vision and Pattern Recognition. [S. l.]: IEEE Press, 2000, 2: 142-149.
- [3] ZHOU HUI-YU, YUAN YUAN, SHI CHUN-MEI. Object tracking using SIFT features and mean shift[J]. Computer Vision and Image Understanding, 2009, 113(3): 345-352.
- [4] SEVILMIS T, BASTAN M GÜDÜKBA Y *et al.* Automatic detection of salient objects and spatial relations in videos for a video database system[J]. Image and Vision Computing, 2008, 26(10): 1384-1396.
- [5] BHATTACHARYYA S, MAULIK U. High-speed target tracking by fuzzy hostility-included segmentation of optical flow field[J]. Applied Soft Computing, 2009, 9: 126-134.
- [6] KOENDERINK J J. The structure of images[J]. Biological Cybernetics, 1984, 50(5): 363-396.
- [7] LINDBERGER T. Scale-space theory: A basic tool for analysing structures at different scales[J]. Journal of Applied Statistics, 1994, 21(2): 224-270.
- [8] 左军毅, 梁彦, 潘泉, 等. 基于多个颜色分布模型的 Camshift 跟踪算法[J]. 自动化学报, 2008, 34(7): 736-742.
- [9] 刘志强, 蔡自兴. 目标跟踪中巴氏系数误差的分析及其消除方法[J]. 计算机学报, 2008, 31(7): 1166-1174.
- [10] 高健, 黄新汉, 彭刚, 等. 一种简化的 SIFT 图像特征点提取算法[J]. 计算机应用研究, 2008, 25(7): 2213-2216.

(上接第 2661 页)

在 $10^{-3} \sim 10^{-2}$ 范围。如阈值取 10^{-2} , 则该金字塔分解应取 4 层, 即 G_0, G_1, G_2, G_3 。相类似, 以其他图像为实验目标模板, Lily 图像金字塔分解应取 3 层, Westconcordorthophoto 图像金字塔分解应取 4 层。这也和很多实际应用中, 图像金字塔分解常采用 3 层或 4 层相一致。

4 结语

在机器视觉应用中, 目标图像与目标模型的匹配校准, 常应用多分辨率图像金字塔。将 CRB 分析应用于多分辨率目标匹配, 可以获得最佳多分辨率图像金字塔分解, 保证匹配精度。而 CRB 对噪声非常敏感, 正比于噪声方差, 使得 CRB 分析中噪声估计尤为重要。在对目标匹配 CRB 分析的基础上, 对多分辨率目标匹配 CRB 中噪声估计方法进行了探讨, 并采用不同阶次有限差分算子构成高通滤波器组, 分别对图像滤波处理, 应用 IQR 方法与中值滤波原理相结合, 估计 CRB 分析中的噪声方差, 避免了对残留轮廓信息的复杂处理, 实验证明了其有效性。所采用的噪声估计方法不仅可用于 CRB 分析, 也可以应用于其他应用中的图像噪声估计。

参考文献:

- [1] ZITOVA B, FLUSSER J. Image registration methods: a survey[J]. Image and Vision Computing, 2003, 21: 977-1000.

- [2] 金勇俊, 李言俊, 张科. 一种透视变换图像金字塔匹配改进算法[J]. 计算机工程与应用, 2007, 43(24): 78-81.
- [3] ROBINSON D, MILANFAR P. Fundamental performance limits in image registration[J]. IEEE Transactions on Image Processing, 2004, 13(9): 1185-1199.
- [4] CAPODIFERRO L, DI CLAUDIO E D, JACOVITTI G, *et al.* Application of local Fisher information analysis to salient points extraction[C]// Proceedings of the IASTED International Conference on Signal Processing, Pattern Recognition, and Applications, SPPRA 2008. Innsbruck: [s. n.], 2008: 13-15.
- [5] LIU CE, SZELISKI R, KANG SING-BING, *et al.* Automatic estimation and removal of noise from a single image[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2008, 30(2): 299-314.
- [6] RANK K, LENDL M, UNBEHAUEN R. Estimation of image noise variance[J]. IEEE Proceedings of Vision, Image and Signal Processing, 1999, 146(2): 80-84.
- [7] PHAM T Q, BEZUIJEN M, van VLIET L J, *et al.* Performance of optimal registration estimators[C/OL]. [2009-03-01]. http://www.ph.tn.tudelft.nl/~lucas/publications/2005/SPIE2005_Orlando_TPLVea.pdf.
- [8] 满家巨, 沈军. 一种图像噪声准确估计方法[J]. 湖南师范大学自然科学学报, 2008, 31(2): 44-47.